

Evaluating the innovation box tax policy instrument in the Netherlands, 2007-13

Citation for published version (APA):

Mohnen, P., Vankan, A., & Verspagen, B. (2017). Evaluating the innovation box tax policy instrument in the Netherlands, 2007-13. *Oxford Review of Economic Policy*, 33(1), 141-156.
<https://doi.org/10.1093/oxrep/grw038>

Document status and date:

Published: 01/01/2017

DOI:

[10.1093/oxrep/grw038](https://doi.org/10.1093/oxrep/grw038)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Evaluating the innovation box tax policy instrument in the Netherlands, 2007–13

Pierre Mohnen,^{*} Arthur Vankan,^{**} and Bart Verspagen^{***}

Abstract: Patent boxes are tax incentive schemes aimed at stimulating research and development (R&D) in firms by providing favourable tax rates to profits that can be linked to a specific immaterial asset, such as a patent. Because these profits are often hard to separate from other firm profits, patent boxes have been argued to be prone to tax shifting of firms, and tax competition between nations, as they might shift the location of the profits without affecting the location of R&D activities. Whether or not this occurs may also depend on conditions set in the patent box. We evaluate the innovation box policy instrument in the Netherlands, which is essentially a patent box, but without the formal requirement of a patent. We ask whether the innovation box has an effect on local R&D investment of the firm ('additionality'), thereby putting the hypothesis of tax shifting to the test. We find that the innovation box indeed has a positive effect on R&D investment, but the average firm that uses the policy tends to use only a part of the tax advantage for extra R&D investment.

Keywords: R&D policy, R&D tax incentives, patent boxes, tax shifting

JEL classification: O38, H26 O52

I. Introduction

Tax incentives for research and development activities are a popular form of innovation policy. [Straathof *et al.* \(2014\)](#) conclude that in 2014 all member states of the EU, except Germany and Estonia, applied a tax instrument to stimulate R&D. These instruments come in many different forms (we discuss a brief taxonomy below). One particular form is the patent box, which provides a special (low) corporate income tax rate for profits that are the results of immaterial assets (patents and other intellectual property rights that result from R&D, such as plant breeders' rights). The idea is that lower taxes on future (expected) profits from R&D will make R&D investments more attractive.

[Griffith *et al.* \(2014\)](#), [Alstadsæter *et al.* \(2015\)](#), and [Evers *et al.* \(2015\)](#) argue that patent boxes stimulate tax competition between (European) nations, possibly leading

* UNU-MERIT and Maastricht University, e-mail: p.mohnen@maastrichtuniversity.nl

** Dialogic, e-mail: vankan@dialogic.nl

*** UNU-MERIT, e-mail: verspagen@merit.unu.edu

We thank Pim den Hertog, Matthijs Janssen, Bas Straathof, and two anonymous reviewers for helpful remarks. The views expressed and any remaining errors are our own.

doi:10.1093/oxrep/grw038

© The Authors 2017. Published by Oxford University Press.

For permissions please e-mail: journals.permissions@oup.com

to tax havens and eroded tax bases. Evers *et al.* (2015) show that in 2014, 12 members of the EU had a patent box measure. Other tax incentive schemes may be less decisive in terms of tax competition between nations, for example because they work on R&D activities undertaken by researchers that are not as easily moved between nations as the profits on which patent boxes operate.

Empirical evidence on the effectiveness of patent boxes, and based on micro data, is scarce. Straathof *et al.* (2014) extensively survey the evaluation studies on R&D tax incentive measures and are unable to include any study on one of the patent box measures in Europe. De Rassenfosse (2015) argues that patent boxes will have little effect on R&D because they target the back end of the innovation process, but he does not conduct an econometric analysis. Alstadsæter *et al.* (2015) do analyse patent box measures, but do not have any micro data. Therefore, their conclusions relate to the aggregate, between-country level and cannot yield any conclusion on the question of whether the patent box schemes effectively incentivize firms to undertake more R&D in the locality where they obtain the tax credit. Yet an answer to this question is an important argument in the debate about tax competition between nations and the potential role that patent boxes play in this process.

We present a micro-econometric analysis of the innovation box tax incentive scheme in the Netherlands, for the period 2007–13. The innovation box started out as a patent box, but was broadened to include R&D activities without a patent or other formal property right in 2010. We estimate the effect of the innovation box in terms of the extra R&D that firms perform (the ‘additionality’ of the policy) as a result of the use of the policy, and compare this to the forgone taxes. This yields a so-called ‘bang-for-the-buck’ (BFTB) measure that summarizes the effectiveness of the policy, and which can be seen as an indicator for the degree of additionality. Our results indicate that the BFTB is positive, i.e. the innovation box effectively stimulates R&D, but is smaller than 1, i.e. the extra R&D is less than the forgone tax income. We discuss the implications of this finding for the debate about tax competition below.

The rest of the paper is organized as follows. In section II, we briefly review some relevant literature. This includes a brief discussion of the special nature of patent box tax measures, and how these affect the probability of tax shifting and tax competition. From this, we also derive our main research question, which is to investigate whether the measured effects of the Dutch innovation box actually indicate that tax shifting is a problem. Section III briefly describes the innovation box policy measure, while section IV introduces our econometric approach. The data is described in section V. Section VI presents the estimation results, including the BFTB for various years and various specifications. The final section VII concludes.

II. A brief literature survey

Straathof *et al.* (2014) present a number of dimensions on which tax incentive schemes for R&D can be distinguished. One dimension is the type of tax that the scheme operates on. This is usually the corporate income tax, but can also be based on social security payments or wage taxes. For example, in the Netherlands, in addition to the innovation

box, firms can also use the so-called WBSO measure,¹ which allows them deduct an amount from the employer's social security contributions on behalf of their employees. This scheme lowers wage costs for R&D.

A second dimension is whether the tax credit works on the total amount of R&D that the firm performs (level-based), or the additional R&D as compared to a base period (incremental). Level-based schemes may have an important deadweight loss, which means that R&D that would have been carried out without the policy scheme is also supported. Increment-based schemes are arguably less sensitive to this, and therefore may be associated with a larger policy effect per euro of forgone taxes. But even in this case the deadweight loss is not zero, as some of the increment may have occurred without the policy.

A final, third dimension is the type of firms that are eligible for the policy instrument. Often, tax credit schemes are generic, i.e. all firms can use them. In other cases, the scheme can be aimed at young firms, or at small and medium-sized firms (SMEs). [Straathof et al. \(2014, p. 21\)](#) conclude that:

[i]n some . . . countries . . . SMEs tend to respond more strongly to the support for R&D, while the reverse was found in other countries . . . [t]here is some evidence that the impact for start-up firms can exceed the average impact, but in general, there is not much evidence on how effectiveness of tax incentives varies with firm age.

R&D tax credits are generally more generous for SMEs because they are more likely to be liquidity constrained, and various studies have shown that SMEs are more responsive in their R&D expenditure to R&D tax credits (e.g. [Hægeland and Møen \(2007\)](#) for Norway; [Corchuelo and Martínez-Ros \(2008\)](#) for Spain; [Lokshin and Mohnen \(2012\)](#) for the Netherlands).

R&D tax measures have been evaluated using a range of econometric methods ([Ientile and Mairesse \(2009\)](#) provide an early survey). One approach is based on the so-called user-cost of R&D (e.g. [Lokshin and Mohnen \(2012\)](#); an early survey is provided by [Hall and Van Reenen \(2000\)](#)). This approach takes R&D as the dependent variable and includes the user cost of R&D (which varies over time and between firms, in particular with respect to the tax incentive scheme) as an independent variable. By identifying the effect of the tax credit on the user costs of R&D, the effect of the policy in terms of extra R&D can be calculated. [Lokshin and Mohnen \(2012\)](#) estimate the effect of the Dutch WBSO, and find a positive BFTB<1. [Mulkay and Mairesse \(2013\)](#) find a similar effect for France in an *ex-ante* evaluation of the 2008 reform, which replaced the incremental R&D with a level-based R&D tax credit.

A different approach is to estimate the effect of the policy in a more direct way, by using a model that distinguishes firms that use the policy from a control group that does not use the policy. This estimates the treatment effect of the policy, for which a range of econometric approaches can be used, such as difference-in-differences, and matching methods (which are also the methods that we use). [Cornet and Vroomen \(2005\)](#) use a

¹ WBSO stands for 'Wet Bevordering Speur- en Ontwikkelingswerk', in full 'Wet Vermindering Afdracht Loonbelasting en Premie Volksverzekering, Onderdeel Speur- en Ontwikkelingswerk' (see [de Jong and Verhoeven, 2007](#)), which translates as 'law for lowering wage taxes and social security contributions related to R&D activities'.

difference-in-differences method to estimate the effect of the Dutch WBSO, and find a positive BFTB < 1. [Duguet \(2012\)](#) uses matching and binary regressions for France, and finds a BFTB approximately equal to 1.

A common problem in the econometric evaluation of the tax credit schemes lies in various kinds of endogeneity. In addition to the endogeneity leading to a selection bias, which we discuss below, there is the effect that the policy may have on R&D costs, in particular wages. When the tax incentive scheme is widely used, it will increase the demand for R&D workers, and hence may drive up the wage rate. This raises R&D expenditures through a price rather than a volume effect. [Lokshin and Mohnen \(2013\)](#) find such a wage effect for the Netherlands and incorporate it in the estimate of the policy effect.

The literature points to the possibility that the specific form of the patent box may lead to tax shifting behaviour of firms, and tax competition between countries, because in this particular form of policy, the tax advantage is linked not to R&D activities but to the profits that result from R&D. As [Straathof et al. \(2014, p. 45\)](#) put it:

Patent boxes . . . offer a large scope for tax planning by firms. They are also part of the strategy of at least some countries to increase tax revenues from foreign companies. Especially for companies with many patents, it can be very difficult to assess what part of a company's income is derived from which patent. This leaves much room for companies and governments to bargain on how much tax has to be paid.

[Alstadsæter et al. \(2015\)](#) argue that this can be influenced by the specific design of the patent box. They point out that some countries require that the R&D underlying the immaterial asset used for the tax advantage is performed locally. Also, some countries rule out the use of bought immaterial assets. Both requirements will tend to limit the possibilities for tax-shifting behaviour by using a patent box. In the Dutch innovation box measure that we evaluate, both these requirements are implemented.

The specific nature of the patent box measures makes them an interesting case for investigation. [Alstadsæter et al. \(2015\)](#) use country-level data to test whether the existence of a patent box policy instrument, and the characteristics of this instrument, are correlated to the level of innovative activity (R&D and patents) in the country. Their conclusions 'confirm [the] fears [that the patent box] could exert a significant effect on patent location without any change in the real research activity, targeting only the tax benefits' (p. 23).

Our micro-econometric analysis of the Dutch case is aimed at finding out whether these fears are also realistic for a case where there is a requirement that the firm developed the immaterial asset used for the tax credit itself. Because of difficulties in assigning profits to a specific R&D project or immaterial asset, this requirement may be ineffective. The evidence that we provide is, to our knowledge, the first evaluation of a patent box measure based on a micro-econometric method.

III. The innovation box

The innovation box measure is a tax instrument that aims to incentivize firms to perform more R&D and to perform this R&D in the Netherlands. Stimulating R&D is the explicit goal of the policy, but because of the emphasis on the returns (profits) to R&D,

this is expected to contribute to innovation and the perceived broad economic benefits of innovation. The instrument works on the corporate income tax, and gives firms a lower rate for profits that are the result of immaterial assets. The measure started (in 2007) as a patent box, i.e. the only immaterial asset that was allowed was a patent. In 2010 it was renamed to innovation box, as the formal patent requirement had been dropped. For profits from immaterial assets, the firm pays 4 per cent tax if its marginal normal rate is 20 per cent, and 5 per cent if its marginal normal rate is 25 per cent. The normal rate is progressive (20 or 25 per cent), and therefore marginal rates differ between firms.

The innovation box is a generic measure that can be used by all firms (also foreign firms active in the Netherlands), but a number of entry criteria apply. The most important one is that the firm must have a patent, a plant breeder's right, or formally recognized (by the tax office) R&D activities. The firm must also generate the immaterial asset itself, i.e. it cannot buy it or have the R&D done externally. There is also a threshold profit level, which is equal to the investment (i.e. the lower rate only applies to profits net of the investment).

Because it is difficult to estimate the exact share of the firm's profits that stems from the immaterial asset, the tax office offers a number of accounting standards to formally determine the profit amount to which the lower rate applies. This can be done on a 'per-asset' basis if the accounting practice of the firm allows for identifying profits from individual assets of R&D projects. Another method is the 'peeling method', where the firm identifies a number of core functions that are essential to the main activities of the firm. A part of the total profit is then allocated to the core functions, based on information on costs. This 'core' profit is then distributed over a number of factors, such as R&D, marketing, production, and entrepreneurship. The part allocated to R&D then applies to the innovation box. The peeling method is applied in cooperation with the tax office. A third method is the cost-plus method, which identifies a standard cost level for the immaterial asset or R&D, and then uses a standard mark-up to determine the innovation box profits.

Several changes to the policy were introduced through the years. The first of these changes took place in 2008, i.e. 1 year after the policy started. Subsequently, changes were introduced every year up to 2013, with the exception of 2012. While the measure started as a patent box, i.e. only applied to projects that led to a patent, this requirement was dropped already in 2008, with the aim to make the policy more interesting for small and medium-sized firms. However, while the limit to the tax advantage for projects with a patent remained at four times the R&D costs, a harder €400,000 limit for projects without a patent was introduced. In 2009, the policy was changed to allow transfer of negative profits of the innovation project between years (both forward and backward in time). In 2010, the effective tax rate was lowered from 10 to 5 per cent, and all limits to the tax advantage were dropped. This mattered most for projects without a patent, as the limit for projects with a patent was almost never reached. Therefore, in effect, this greatly enhanced the opportunities for use of the policy without a patent. In the same year, the name was changed to innovation box. In 2011, profits that occurred after a patent was applied for but before the patent was granted were allowed to be included in the tax advantage. Finally, in 2013, a fourth and much simplified method was allowed to calculate the innovation profits, in addition to the three methods discussed above. Altogether, these changes imply that the policy has become more accessible to a larger group of firms over time.

IV. A short introduction to the methods

A key factor in deciding about which method to use is the endogeneity of the innovation box tax scheme. Whether or not a firm uses the policy is a voluntary decision. Firms who decide to use the policy will likely make this decision because of certain factors that are not shared with firms that do not use the policy. An observed increase in R&D for the users of the policy may well be the result of these specific factors, rather than of the policy itself. This is evident for the basic distinction between innovators and non-innovators: non-innovators will be less likely to use the policy. If we then compared policy-users and non-users in a naïve way, we could conclude that the policy works because users have a larger increase in R&D than non-users. But this conclusion would be wrong, because the differences between innovators and non-innovators already existed before the policy was introduced.

From the point of view of evaluation, it would be better if the use of the policy were determined in an exogenous way, e.g. in a lottery. In this way, differences between users and non-users of the policy could be made essentially random, so that it would be likely that actually observed differences in increase of R&D are the result of the policy. However, because the innovation box is not implemented in this way, we have to resort to econometric methods that control as much as possible for the endogeneity, but still enable us to draw conclusions about the effect of the policy. We apply a difference-in-differences method, which corrects as much as possible for selection bias and endogeneity.

We always use a balanced panel of firms, i.e. all firms in the analysis are available for the entire time period of the estimation. This reduces the number of observations, but gives higher reliability, because in an unbalanced panel the number of firms in the control group (non-users) would fluctuate strongly, which means that the users of the policy are compared to different firms for every period (year).

The difference-in-differences method is typically implemented in a situation where observations exist for two periods, one with and one without the policy. In this case, the difference-in-differences is implemented by a regression of the following kind, for a sample of firms:

$$Y_{it} = b_0 + b_1 D^1 + b_2 B_i + b_3 D^1 B_i + \sum_j \beta^j X_{it}^j + \varepsilon_{it}. \quad (1)$$

Here, Y is the dependent variable by which we measure the effect of the policy (R&D), and X^j is a range of control variables. D^1 is a dummy-variable equal to 1 for period 1 (when the policy exists) and 0 for period 0 (when the policy does not exist and hence no firm uses the policy). B is a dummy variable equal to 1 for firms that use the policy (in period 1) and 0 for the control group of firms that do not use the policy (in either period). The estimated parameter b_0 measures the average outcome (in terms of Y) for the control group in period 0. $b_0 + b_1$ is the average outcome for the control group in period 1. b_2 is the difference between the control group and the users in period 0. This difference already exists before the policy is implemented, and is therefore not a part of the effect of the policy. $b_0 + b_2$ is the average outcome for users of the policy in period 0, while $b_0 + b_1 + b_2 + b_3$ is the average outcome for these firms in period 1. The difference in outcome for users between period 0 and period 1 is therefore $b_1 + b_3$. Of this, b_1 is identical to the difference for the control group. Hence, b_3 , which is the

difference-in-differences, measures the ‘additionality’ of the policy, which is the extra R&D performed as a result of the use of the policy.

In our case, data are available for more than two periods, which is why we implement a difference-in-differences method with more than one period of policy availability. There are two ways in which equation (1) can be transformed to such a multiple period case:

$$Y_{it} = \sum_{T \neq 0} \gamma^T G_i^T + \sum_T \tau^T D^T + \sum_{T_0} \sum_{T_1 \geq T_0} \alpha^{T_0 T_1} G_i^{T_0} D^{T_1} g_{it} + \sum_j \beta^j X_{it}^j + \varepsilon_{it}, \quad (2a)$$

$$Y_{it} = \sum_{T \neq 0} \gamma^T G_i^T + \sum_T \tau^T D^T + \sum_{T_0} \sum_{T_1 \geq T_0} \alpha^{T_0 T_1} G_i^{T_0} D^{T_1} + \sum_j \beta^j X_{it}^j + \varepsilon_{it}. \quad (2b)$$

G^T is a dummy variable equal to 1 for a firm that starts using the policy in period T (0 remains the period without the policy, hence T can be equal to 1 or larger). D^T is a dummy variable for time period T , while g_{it} is a dummy variable that indicates whether firm i uses the policy in period t .² T_0 represents the period just before the first use of the policy, and T_1 any other period after this. The parameters γ , τ , α and β are estimated. The γ parameters correct for differences between users and non-users that already existed before the introduction of the policy, to the extent that these differences are not reflected in the set of variables X . We use multiple γ parameters to enable a distinction between different categories of users, such as early and late adopters of the policy (non-users will have zero value for all G variables, hence no γ refers to them). In other words, we assume that firms that begin using the policy immediately after it has been introduced will differ from firms that start using the policy only after a few years. Because, as discussed in the previous section, several changes in the policy occurred, it is advisable to account for these differences. The α parameters measure the effect of the innovation box policy. Instead of just a single effect, we estimate one effect for each combination of user-generation (G) and time period. For example, the parameter α^{0812} measures the effect of the policy in 2012 on firms from the 2008 generation. A similar parameter (effect) is estimated for every possible combination of period and generation.

In the firm’s decision about whether to increase R&D efforts as a result of the policy, several time lags may play a role. The profits that are relevant for the innovation box will generally occur after the innovation efforts took place. A completely rational firm will therefore make an estimate of the expected profits, the time lag at which they occur, and then compare the benefits including the innovation box with the innovation costs that have to be invested up front. In such a case, innovation box benefits in the future will affect current innovation efforts. However, because the innovation process is associated to strong uncertainties, it is also possible that firms decide in a different way, with a less strict role for the time lag between investment and pay-off. For example, the firm may use the innovation box benefit to contribute to innovation efforts in the current year.

Such ambiguities in the time structure of the effect lead us to experiment with two variations in the estimations below. First, we adopt the two alternative specifications (2a) and (2b). The difference between these specifications lies in the assumption about how long the effect of the innovation box policy lasts. Equation (2a) assumes that the

² As a matter of notation, T is used for variable names whereas t is used to indicate time periods.

effect is limited to the time period in which the use of the policy occurs (of course the firm may use the policy in different years, so that an effect also occurs in multiple years). On the other hand, equation (2b) assumes that the firm will always have an effect as a result of a one-period use of the policy, even if later it does not use the innovation box. The two specifications would be identical if first-time users continued using the policy forever. Because there is a large persistency level in innovation box use, reality is actually close to such a situation.

Second, we introduce the so-called anticipation effect. This is inspired by the idea that expected future profits determine current R&D efforts, and hence that the use of the innovation box in period t can already affect R&D efforts in periods before t . As an approximation, the anticipation effect assumes that this time lag is exactly 1 year. We implement this by shifting the policy use (dummy) variable 1 year back in time, e.g. when a firm actually uses the innovation box in 2010, the anticipation effect specification attributes the use to the year 2009. Note that if the anticipation effect is relevant, but is not taken into account, the result may be that the actual effect of the policy is interpreted as an *ex ante* difference between users and non-users. We do not explore an anticipation effect longer than 1 year because this would eliminate almost a third of our observations. The β parameters articulate the differences between firms that are observable in terms of the variable set X . The τ parameters measure how the dependent variable varies over time, on average, and independent of the policy. Because the credit crisis of 2007 and the slow economic recovery following it are part of our analysis period, this is an important control variable.

We do not include spillover effects in the estimations, even if it is well known in the literature (e.g. Griliches, 1992) that (additional) R&D by one firm can stimulate other firms to increase their R&D expenditure (assuming strategic complementarity), both within the same industry and in other industries. In fact, we assume that the spillovers affect all firms in the sample, and hence are not a part of the measured policy effect. We also do not attempt to estimate the policy-induced spillovers, i.e. the spillovers of the R&D generated by the policy, which will lead to an under-estimation of the estimated policy effect.

V. Data

We use a dataset that results from the other R&D tax credit measure in the Netherlands, WBSO (in later years also RDA, which is similar).³ For WBSO/RDA, firms have to supply information on the number of R&D person-hours on a yearly basis. For the period under analysis here, virtually all firms use WBSO/RDA, and hence the dataset provides an almost complete population of enterprises that do any kind of R&D in the Netherlands. This is in strong contrast to another potential dataset for our analysis, the Community Innovation Survey (CIS) as implemented by Statistics Netherlands. CIS is an often-used data source for innovation analysis, but it is based on a sample of firms, with samples being small especially for small and medium-sized firms. We tried

³ As already explained, WBSO reduces the wage costs of R&D workers, whereas RDA are tax credits for other R&D costs than labour costs.

matching the CIS data with data on innovation box use, but were only able to match a small number of users. Therefore, we only present results for the WBSO dataset.

Data on the use of the innovation box policy is defined at the level of so-called tax-units, which are legal entities used by the tax authorities for collection of the corporate income tax. The tax-units can be linked on a one-to-one basis to the statistical units used by Statistics Netherlands, which is the supplier of the data used in this study. However, neither useful economic data nor information about the use of the innovation box measure are available at these levels of tax-units and statistical units. Information about innovation box use comes at the level of tax entities, which are aggregations of tax-units. Statistical information comes at the level of enterprises, which are aggregations of statistical units. Although they are composed of the same kind of units, statistical enterprises are generally not identical to a tax entity.

In order to construct our dataset, we need to link the statistical enterprises to tax entities. Because these entities do not correspond to each other on a one-to-one basis, we need to make compromises in linking the data. For about half (52 per cent) of all cases, we are able to make this link in a ‘pure’ way: there is a full coverage of all the units of the tax entities and all those of the statistical enterprises, hence a one-to-one match between tax entities and statistical enterprises. For 40 per cent of the remaining cases, there is only one tax entity for each statistical enterprise, but some of its units also belong to another statistical enterprise, i.e. there is some overlap between tax entities and statistical enterprises.

Table 1 gives summary statistics for the dependent variable (R&D person-hours). It is immediately seen that innovation box users have much higher R&D efforts than non-users. Interestingly, however, the number of R&D person-hours declines over time within the group of users, as well as in the group of non-users. In the user-group, this effect is especially large in the beginning, which indicates that the use of the innovation box in the early years was mainly a matter of a small group of very large firms.

Table 1: Summary statistics for the dependent variable (R&D person-hours)

Year	Number of firms			Average person-hours for R&D		
	Non-users	Users	Total	Non-users	Users	Total
2007	10,913	40	10,953	6,696 (49,066)	105,377 (413,677)	7,056 (55,168)
2008	11,154	236	11,390	6,511 (51,169)	24,902 (118,867)	6,892 (53,501)
2009	13,598	436	14,034	5,937 (45,209)	21,262 (101,649)	6,414 (48,039)
2010	13,578	929	14,507	5,604 (36,531)	21,989 (119,189)	6,654 (46,626)
2011	14,116	1,394	15,510	6,103 (55,462)	20,356 (112,912)	7,384 (62,939)
2012	15,872	1,742	17,614	4,589 (35,907)	20,445 (120,956)	6,157 (51,287)
2013	14,856	1,894	16,750	5,349 (39,356)	13,258 (80,866)	6,243 (46,034)

Source: Statistics Netherlands microdata, processed by Dialogic/MERIT in September 2015. Numbers in parentheses are standard deviations.

VI. Results

We use the natural logarithm of the number of person-hours spent on R&D as the dependent variable in the difference-in-differences estimations. We have no data for the year 2006, while the innovation box was introduced in 2007. Thus, strictly speaking, we have no data on the year preceding the introduction of the policy. We deal with this by eliminating from the dataset all firms that used the innovation box in 2007, thus effectively making 2007 the reference year in which no firm used the innovation box.⁴ Since some changes to the patent (respectively innovation) box were introduced every year, we do multiple comparisons of difference-in-differences, namely we compare the users of every year from 2008 onwards to the firms that never used the patent box from 2007 to 2013. We have (the natural logarithm of) the number of employees of the firm and a set of sector dummies as the control variables X in equations (2a) and (2b).

Table 2 documents the estimated parameters for the policy effects. Other estimated parameters are not documented in order to save space. The two columns refer to specification (2a), i.e. assuming the policy has a cumulative effect. In the first column we estimate the model under the assumption of no anticipation effect. This column finds a positive and significant effect for just about half of the combinations of user-generations and time periods. This indicates that for these periods, the specified user-generations actually carried out more R&D than the control group, or, in other words, that the policy was effective. The large majority of these cases is found in the years 2012 and 2013. Before 2012, the innovation box seems not so effective. The size of the effects is fairly large. Because the dependent variable is in natural logs, the documented effects can approximately be interpreted as percentage effects. The next column gives the results for the anticipation effect specification. This column shows 12 significant positive effects, one fewer than without an anticipation effect. Otherwise, there are no strong indications that the models with or without anticipation effects yield different results. We have also estimated the model under specification (2a) and no anticipation.⁵ In that case we obtain only eight significant coefficients.

We also experimented with propensity score matching as an alternative evaluation method. However, since we have few control variables that we can rely on to construct the counterfactuals, it is difficult to dissociate the endogenous selection from the true effect of the patent/innovation box. The propensity score method yields higher effects, but they are for instance smaller already when we restrict the control group to firms that at least once over the sample period have used the patent/innovation box. We have more confidence in the method of difference-in-differences and continue the analysis with the results reported in Table 2.

In order to get a better insight into the size of the policy effects, we summarize the most important results in a graphical way. The first step is to transparently document the benchmark against which the innovation box use is evaluated. This is done in Figure 1, which gives the predicted number of R&D person-hours for the different generations of innovation box users, and non-users, over time. The figure assumes that the innovation box never existed (i.e. the α parameters are set to zero) and uses average firm size and no sector-effects. It is based on the first column of Table 2.

⁴ Only 40 firms out of 10,953 made use of the patent box in 2007.

⁵ The fact that so few firms used the patent box at the beginning (see Table 1) shows that there is little sign of an anticipation effect.

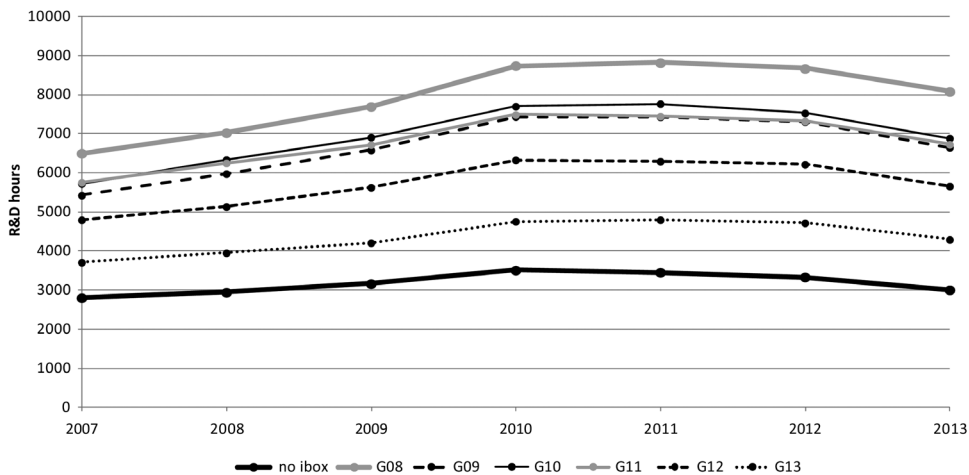
The lines in the figure are the sum of a yearly effect (equal to all groups of firms), a generation effect (G , non-existent for the non-user category), and a size effect (for every group, the yearly average within the group was used). The resulting differences between the groups of firms are sizeable, which means that it is important to eliminate

Table 2: Difference-in-differences estimation results (only policy effects, by year and generation of users)

Generation of users (G)	Year (t)	(1)	(2)
		No anticipation	Anticipation
2008	2008	0.021 (0.055)	0.032 (0.046)
2008	2009	-0.002 (0.064)	0.128 (0.059**)
2008	2010	0.002 (0.069)	0.083 (0.059)
2008	2011	0.052 (0.067)	0.115 (0.071)
2008	2012	0.118 (0.070*)	0.186 (0.072***)
2008	2013	0.203 (0.075***)	0.198 (0.086**)
2009	2009	0.112 (0.049**)	-0.008 (0.034)
2009	2010	0.065 (0.055)	0.021 (0.038)
2009	2011	0.092 (0.067)	0.064 (0.043)
2009	2012	0.161 (0.067***)	0.139 (0.047***)
2009	2013	0.179 (0.080**)	0.208 (0.053***)
2010	2010	0.022 (0.032)	0.016 (0.031)
2010	2011	0.061 (0.038)	0.010 (0.038)
2010	2012	0.135 (0.042***)	0.075 (0.042*)
2010	2013	0.208 (0.050***)	0.159 (0.043***)
2011	2011	-0.000 (0.032)	0.127 (0.040***)
2011	2012	0.063 (0.037*)	0.153 (0.051***)
2011	2013	0.153 (0.039***)	0.273 (0.056***)
2012	2012	0.121 (0.045***)	0.106 (0.047**)
2012	2013	0.246 (0.051***)	0.154 (0.058***)
2013	2013	0.137 (0.053***)	—

Notes: Numbers in parentheses are p -values; one, two, and three stars indicate significance at 1, 5, and 10 per cent levels, respectively.

Figure 1: Estimated differences between generations of users and non-users of the innovation box policy

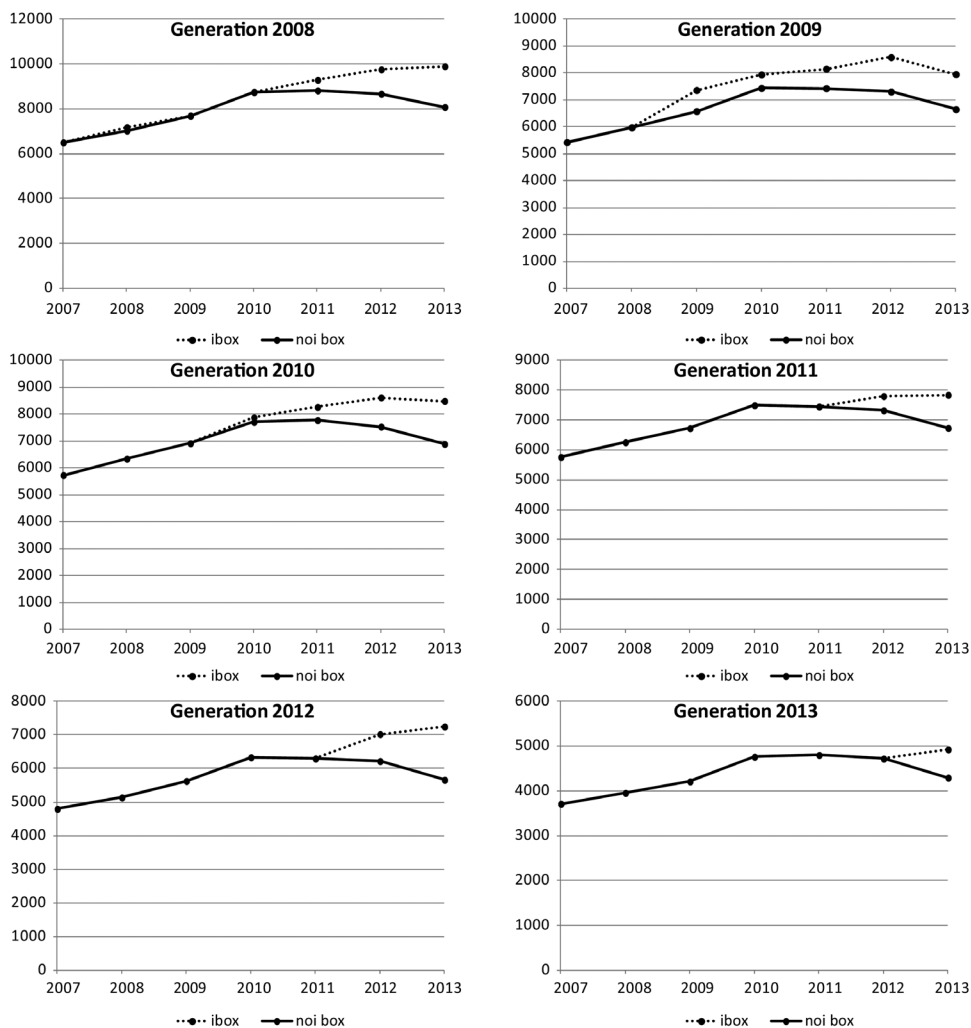


these differences from the actual effect of the policy. The group of non-users has the lowest R&D effort over the entire period. Thus, simply comparing non-users and users in order to estimate the policy effect gives a strongly distorted picture. Such differences are also observed between the different generations of users. The first generation (G08) has the highest R&D effort. The next three generations (G09, G10, and G11) are not much different, but the last two generations (G12 and G13) exert clearly still a lower R&D effort even in the absence of the patent box than the previous three generations of first users. So it is clear from Figure 2 that the higher the R&D effort, the earlier the firms make use of the patent box. This is broadly consistent with our review of changes in the innovation box policy in section III.

Figure 2 (also based on the first column of Table 2) shows the most elementary representation of the policy effects. Each sub-figure compares the R&D effort inclusive of the policy effect of an average user firm with the counter-factual of no innovation box use. The latter benchmark is taken from Figure 1. As we see, every successive generation of first users starts from a lower benchmark, but follows a similar counterfactual evolution of R&D person-hours in the absence of the patent/innovation box policy. The 2008 generation of first users does not change its R&D efforts until the year 2010, when the innovation box was introduced. The policy effect is most clearly visible for the generation of 2009, where the dotted line is always above the solid line. This is also the case for later generations, but here there are fewer years with innovation box use available. In general, the figures clearly show the nature of the 2010 extension of the patent box to the innovation box: it is mostly a compensation of a downward trend that shows in the benchmark after 2010. It is also seen that the effect is often substantial in terms of the number of R&D person-hours.

In order to better understand the policy effects, we calculate the so-called bang-for-the-buck (BFTB) measure. This terminology is commonly used in the Dutch policy debate to express the effect of R&D incentive policies in terms of additional R&D as a fraction of the forgone tax income for the government. A BFTB value of 1 would imply that for every euro of forgone tax income, an additional euro of R&D is undertaken by

Figure 2: Innovation box effects on different user generations against the benchmark of no use, difference-in-differences estimations (R&D hours on the vertical axis of each figure)



the firm. BFTB lower (higher) than 1 indicates that less (more) R&D is generated by the policy than forgone tax income.

In our dataset, we have no data on R&D expenditures, only on R&D person-hours. In order to convert R&D person-hours to R&D expenditure, we first estimate a regression model for R&D expenditures using the CIS dataset. The model includes as explanatory variables R&D person-hours and all control variables used in the difference-in-differences model, which means that we can ‘predict’ R&D out-of-sample, i.e. for all firms in the dataset used for Table 2. In this model, we also include interaction effects between the independent variables in order to maximize the fit. The (out-of-sample) predictions are used to calculate the BFTB in the following way. The extra R&D person-hours due to the innovation box are estimated, per firm, using the estimates of Table 2 and then

expressed as a percentage of the total estimated R&D person-hours of the firm in the same model (Table 2). This percentage is applied to the out-of-sample predicted R&D expenditure of the firm in the same year, which gives us the ‘bang’ in the BFTB. We use the firm’s own generational effect, which implies that the total sample ‘bang’ is a weighted average for all user generations. The ‘buck’ is the innovation box tax deduction in the same year (in case of the anticipation effect, the ‘buck’ is also advanced by 1 year). The average BFTB that will be reported is obtained by dividing the ‘bang’ by the ‘buck’.

The BFTB values in Table 3 range from 0.07 to 0.84. The higher values, 0.73 and 0.84, are found for 2012, which is consistent with the previous impression that the strongest effect is found for the later years (for 2013, no R&D expenditure is available, so no BFTB can be estimated). In 2009, the model without anticipation effect also estimates a high BFTB, and 2011 yields relatively high values for both estimates. The brief summary of changes of the measure that we presented in section III suggests that the policy becomes more accessible for a broad population of firms (e.g. because the requirement of a patent is dropped). The relatively high BFTB values for the two most recent years may indicate that this leads to a stronger effect of the policy, although the changes over time in Table 3 remain somewhat erratic, and therefore do not warrant a very strong conclusion.

All values in the table are below 1. This means that the extra R&D generated by the innovation box policy is less than the forgone tax income. In other words, firms do not transform every euro of tax credit into extra R&D. The model specifications with anticipation effect do not yield qualitatively different results than without anticipation. The average values for 2008–12 are about one-half (without anticipation) or one-third (with anticipation).

A BFTB lower than 1 is the most common result in the literature, although some studies find values (much) above 1. The summary table in Straathof *et al.* (2014, p. 33) documents 10 estimated values for the BFTB for a range of countries and time periods. Of these 10 values, four are equal to 1 or larger, and the remaining six are positive but smaller than 1.⁶ For the Netherlands, a number of studies have documented the BFTB for the WBSO instrument. Two studies document BFTB>1: de Jong and Verhoeven (2007) report BFTB equal to 1.72, and Verhoeven *et al.* (2012) find 1.77. Two other studies find BFTB<1: Cornet and Vroomen (2005) report BFTB equal to 0.15 (general) or 0.65 (start-ups), and Lokshin and Mohnen (2012) find 0.54.

Table 3: Estimated BFTB for innovation box

Year	No anticipation effect	Anticipation effect
2008	0.21	0.48
2009	0.70	0.08
2010#	0.22	0.07
2011	0.43	0.36
2012	0.73	0.84
2008–12	0.54	0.34

Notes: For calculation method see main text. # In 2010 a different model was used to estimate R&D expenditures (less interaction terms; the standard model yielded many negative estimates).

⁶ The magnitude of the BFTB depends on the way it is calculated and on the type of R&D tax incentive. For incremental tax incentives it is by construction greater than 1.

The fact that the BFTG is found to be lower than 1 can be explained in the following way. The patent/innovation box is an *ex post* reward for successful R&D. It can stimulate firms to do more R&D, but it is not warranted that the R&D will lead to patents (in the case of the patent box) or be eligible for the patent or the innovation box because of restrictions attached to it (the R&D must have been declared previously, the R&D must be done at home . . .). Moreover, it is not straightforward to isolate the revenues that can benefit from the innovation box. Therefore the stimulus to do more R&D is probably lower than with an equal amount of *ex ante* tax incentives, the more so as the return to R&D in terms of patents or other revenues is expected to be low.

VII. Discussion and conclusions

We performed a micro-econometric evaluation of the Dutch innovation box policy instrument. These instruments, generally known as ‘patent boxes’, have been argued to lead to tax-shifting behaviour and tax competition between countries. Our evaluation is the first available one based on a micro analysis. The available macro evidence (Alstadsæter *et al.*, 2015) suggests that tax-shifting behaviour does occur in the case of patent boxes, although the specific design of the instrument may counteract this. In particular, the requirement that the tax advantage must be linked to the firm’s own (therefore local) R&D activity may mitigate the role of patent boxes in tax shifting. The Dutch case that we analysed has such a local content requirement, and therefore we can look at our case as a test of whether such requirements can be effectively implemented.

We find that the innovation box has a positive effect. Firms that use the policy instrument tend to have higher R&D activities after they start using the innovation box. This result is obtained in a difference-in-differences estimation, which represents the best possible way that we can account for endogeneity of the policy use and the selection bias that is associated to this. A sensitivity analysis using a different method (propensity score matching) indeed shows that the choice of method matters, as we obtain stronger positive effects with the propensity score matching. However, we prefer the difference-in-differences estimations because they are *a priori* expected to yield a purer cleaning of the selection bias.

We calculate a ‘bang-for-the-buck’ (BFTB) that measures the extra R&D generated by the policy instrument per euro of forgone taxes. The BFTB that we find varies over the years, but is generally positive and smaller than 1. Hence we conclude that there is additionality of the innovation box measure, but only at a modest level. Although the measure is effective in generating extra R&D, it does so to a lesser extent than the costs that are involved in terms of forgone taxes. The patent/innovation box is thus an effective measure to stimulate additional domestic R&D, but it is not efficient in terms of cost–benefit analysis. A proper cost–benefit analysis would also include the administration and application costs as well as the deadweight loss from taxation on the cost side, and the positive R&D externalities on the benefit side. Our expectation is that the latter would have to be substantial to overcome the various costs and turn the measure to be welfare improving. We leave this for further research. But, on the other hand, the BFTB values that we obtain are broadly in line with those obtained for tax credit policy schemes of a different nature than a patent box, in a range of countries. This suggests that at least the Dutch innovation box instrument is not less effective than tax

credit instruments that are not of a patent box nature. We therefore conclude that it is indeed very well possible ‘to mitigate the role of patent boxes as new tax competition tools’ (Alstadsæter *et al.*, 2015, p. 23). At least when implemented as in the Netherlands, the patent/innovation box policy is a way to limit the R&D tax incentive to domestic R&D, to include R&D that is not patent-bound, and yet to restrict the R&D support to successful R&D (which is not the case with *ex ante* tax incentives). It would be interesting to compare, in a full cost–benefit analysis, the *ex post* innovation box tax incentive with the more traditional level-based *ex ante* tax incentive measure (WBSO in the Netherlands) and to find out whether the presence of both measures within a policy mix is justified.

References

- Alstadsæter, A., Barrios, S., Nicodeme, G., Maria Skonieczna, A., and Vezzani, A. (2015), ‘Patent Boxes Design, Patents Location and Local R&D’, *Cesifo Working Paper* No. 5416.
- Corchuelo, B., and Martínez-Ros, E. (2008), ‘Aplicación de los Incentivos Fiscales a la Inversión I+D en las Impresas Españolas’, *Hacienda Pública Española/Revista de Economía Pública*, **187**(4), 9–39.
- Cornet, M., and Vroemen, B. (2005), ‘Hoe effectief is extra fiscale stimulering van speur- en ontwikkelingswerk? Effectmeting op basis van de natuurlijk-experimentmethode’, CPB Document 103, The Hague, Central Plan Bureau (CPB).
- De Jong, J. P. J., and Verhoeven, W. H. J. (2007), *Evaluatie WBSO 2001–2005: Effecten, Doelgroepbereik en Uitvoering*, Ministry of Economic Affairs, the Netherlands.
- De Rassenfosse, G. (2015), *Patent Box Policies*, Office of the Chief Economist, Commonwealth of Australia.
- Duguet, E. (2012), ‘The Effect of the Incremental R&D Tax Credit on the Private Funding of R&D: An Econometric Evaluation on French Firm Level Data’, *Revue d’Economie Politique*, **122**(3), 405–35.
- Evers, L., Miller, H., and Spengel, C. (2015), ‘Intellectual Property Box Regimes: Effective Tax Rates and Tax Policy Considerations’, *International Tax and Public Finance*, **22**, 502–30.
- Griffith, R., Miller, H., and O’Connell, M. (2014), ‘Ownership of Intellectual Property and Corporate Taxation’, *Journal of Public Economics*, **112**, 12–23.
- Griliches, Z. (1992), ‘The Search for R&D Spillovers’, *Scandinavian Journal of Economics*, **94**, S29–47.
- Hægeland, T., and Moen, J. (2007), ‘Input Additionality in the Norwegian R&D Tax Credit Scheme’, reports 2007/47, Statistics Norway.
- Hall, B., and Van Reenen, J. (2000), ‘How Effective Are Fiscal Incentives for R&D? A Review of the Evidence’, *Research Policy*, **29**, 449–69.
- Ientile, D., and Mairesse, J. (2009), ‘A Policy to Boost the R&D: Does the Tax Credit Work?’, *European Investment Bank Paper* **14**(1).
- Lokshin, B., and Mohnen, P. (2012), ‘How Effective are Level-based R&D Tax Credits? Evidence from the Netherlands’, *Applied Economics*, **44**, 1527–38.
- (2013), ‘Do R&D Tax Incentives Lead to Higher Wages for R&D Workers? Evidence from the Netherlands’, *Research Policy*, **42**, 823–30.
- Mulkay, B., and Mairesse, J. (2013), ‘The R&D Tax Credit in France: Assessment and Ex Ante Evaluation of the 2008 Reform’, *Oxford Economic Papers*, **65**, 746–66.
- Straathof, B., Gaillard-Ladinska, E., Kox, H., Mocking, R., Goldberg, I., Jensen, C., Lindholm, P., Sobolewski, M., Berg-Andersson, B., Kaseva, H., Rouvinen, P., Valkonen, T., Sellner, R., Schönplflug, K., and Paterson, I. (2014), ‘A Study on R&D Tax Incentives’, Taxation papers, Taxation and Customs Union Working Paper n. 52–2014, Brussels, European Commission.
- Verhoeven, W. H. J., Van Stel, A. J., and Timmermans, N. G. L. (2012), *Evaluatie WBSO 2006–2010*, Ministry of Economic Affairs, the Netherlands.